Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2017, Vol. 9, No. 3

Processing Scalar Implicatures in Mandarin Chinese: Testing the Processing Models

Si Liu

The School of Foreign Languages and Literature, Lanzhou University, Lanzhou, Gansu, China

E-mail: zsmyjjk@126.com

Jianan Liu

Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, Netherlands

E-mail: j.liu@students.uu.nl

Received: March 26, 2017 Accepted: April 8, 2017 Published: June 22, 2017

doi:10.5296/ijl.v9i3.11432 URL: https://doi.org/10.5296/ijl.v9i3.11432

Abstract

This study first adopted a participant-perception test to assess the processing model of scalar implicature in Chinese. Our main aim is to distinguish among the three possible processing mechanisms: the context-driven account, the default account and the standardized account. We designed two experiments to testify these three models mentioned above: one without any context and the other one with upper and lower contexts. In our Experiment 1, we conducted test items without contexts in child and adult groups, whose aim is to test the necessity of context to scalar implicature and thus clearly discern the three models. We found though without context, both children and adults group processed scalar implicature at a medium rate, which was an evidence to deny the context driven account. However, some adults tended to be confused about the experiment purpose when facing testing items totally without context constraints, and the children participants might get help from other developed linguistic ability in their processing, like the improved numeral ability. Thus it would be clear that the context account is unreliable, but it would still be early to tell whether the results support the default account or the standardization. In Experiment 2, we added the context constraints, the upper bound context and the lower bound context. Our final results, the still processing of utterance with SI in lower bound context and a similar reaction time to the scalar implicature processing in both upper and lower bound contexts denied the default

account and showed a closer relation to the standardization account.

Keywords: Pragmatics, GCI, Scalar implicature, Experimental pragmatics, Processing model, Meaning processing, Mandarin Chinese, Children cognition

1. Introduction

1.1 Scalar Implicatures

In our daily conversations, communication is not confined just to its literal meaning. What people mean is not only shown by the words they say, but by the meanings that are implicated. This kind of implicated meaning, which is known as the implicature, arises according to the Cooperative Principle and its four maxims put forward by Grice (1975). One of the maxims, the maxim of Quantity, proposes that when people are cooperative in the conversation, both the addresser and the addressee give as much information as the conversation needs and also no more than the knowledge they have at that moment. Look at Example (1) as follows:

(1) The boss: "Did all the employees attend the meeting this morning?"

The secretary: "Some of them did."

The secretary could have responded to the boss with "all", but that would have been more informative than what she really knew. In order to say the truth and say no more than what she knew, the secretary responded with "some", which holds the implicature that "not all" the employees attended the meeting this morning. Generally, the words "some" and "all" form a scale <some, all>, which is derived from a bigger scale <some, many, most, all>. In another word, the word on the right of the scale is more restraining and thus more informative and stronger than the word on the left. Besides, the affirmation of the weaker item implies to deny the stronger items. In this way, using "some" implies "not many", "not most" and also "not all", specifically, <some, all> composes a scale in which the two words are of the same semantic field but have different informative strength. The word "all" is more constraining than the word "some". In this way, "all" is more informative and stronger, while "some" is less informative and weaker, just as in the example (1) above. Such a kind of implicature as using "some" implied "not all" is called scalar implicature, which was first proposed by Horn (1972). It is only restricted in the semantic facet, that is, all the scale items form a scale based on their semantic properties and entailment relations.

The same situation occurs in other languages except English. A Chinese example of scala rimplicature can also be illustrated in the same way as follows.

Yige guke zoujin yijia shangdian.

a (quantifier) costumer walked into a (quantifier) sstore

A costumer walked into a store.

Zai liulan yibian hou, ta dui yingyeyuan shuo, " Wo xihuan zheli de yixie yifu."

after looking a(quantifier) she to salesgirl said I like here of some clothes

After looking around, she said to the salesgirl, "I like some of the clothes here"

In this example, the salesgirl should have received information from the costumer that she did

not like all of the clothes here in the shop, for the application of "*yixie* (some)" implies the negation of "suoyou (all)".

Scalar implicature is of the most primitive categories and has been widely studied by linguists. Except for the seminal discussion of Grice and Horn, other linguists also have discussed a lot concerning this topic, which could be divided into the Neo- Gricean and Post- Gricean accounts. Liu Si & Liu Huangmei (2014) has made a detailed review about the recent theoretical developments.

1.2 The Scalar Processing Models and Related Experiments

Now, our discussion is focused on the three scalar processing modes: the default account model, the context-driven account model and the standardization account model.

1.2.1 The Default Account

According to Neo-Gricean view, Scalar Implicature SI belongs to the Generalized Conversational Implicature (GCI), which can be generated without dependence on contexts or situations. The default account proposed by Levinson (2000) holds that SI can be derived merely through the literal form in which the processing of an utterance with SI occurs without contexts. In terms of the default mechanism, SI underlies the utterance would be processed immediately by the hearer or reader, due to the processing is based on the literal form of the scalar item. However, this default account does not mean that the SI would always stay away from the influence of context. If the processed SI is not suitable or cannot match with the current contexts, then there would be a cancellation to the SI.

In this way, the presumed psychological reaction concerning default account would be in two aspects. When the context is appropriate, the derivation of SI would be very rapid, which is due to the instant reaction to the literal form of the scalar item. While the context is not appropriate, the reaction time of processing SI utterance would be much slower, which is due to the further cancellation of the already generated SI.

1.2.2 The Context Driven Account

The context driven account that is proposed by Relevance theorists stresses the importance of the context in the processing of utterance (Carston, 1991;2002, Sperber & Wilson, 2001). It proposes that SI is involved in the particularized conversational implicature, whose cognitive process should be accounted with the Relevance Theory. The Relevance theorists hold the view that "what is said" (that is *explicature* in their term) is merely a kind of stimulus from the utterance. The processing of implicature has to rely on the relevance from the context. More specifically, when the hearer or reader encounters a scalar item, he or she cannot process SI literally. The crucial factor determining the process of SI is the context which needs to be relevant enough for the hearer or the reader to process SI. Unlike the default account, context-driven account holds that no cancellation of SI occurs when a context does not exist or the context is not suitable, because SI would not be generated at all in such kinds of situations.

In this way, the presumed psychological reaction to SI derivation would be that in an

appropriate context, the processing of SI utterance could be more time consuming than that in terms of the default account, since the processing would be affected by the context. If no context exists or suits the generation, no SI happens.

1.2.3 The Standardization Account

The standardization theory proposed by Bach and Harnish is first applied in the field of speech act, especially the illocutionary act. Later on, it was also used to interpret GCI. According to the standardization account, as a part of the illocutionary speech act, GCI does not need to be processed every time since it can be standardized through frequent uses. The standardization is part of participant's linguistic knowledge, which includes the meaning of certain structures and also the usage of certain structure in a sentence. In this way, standardization can work in the processing of utterance with SI merely with the literal meaning. This priority can also happen when taking the context into consideration, which is also the most salient characteristic of standardization.

According to standardization, the hypothesis of the psychological reaction to SI would be that the generation of SI would be rapid and less time costing, which would be sustained in any contexts. Even when a context is not suitable, SI would be generated, because it has already been standardized in speakers' minds to the literal structure due to over uses.

1.3 Related Studies

1.3.1 Some earlier studies

Bezuidenhout and Cutting (2002), as well as Bezuidenhaut and Morris (2004), asserted that a pragmatic context driven account is the preferred processing model. Noveck and Posada (2003), as well as Bott and Noveck (2004), also supported the pragmatic model, based on their time-costing experimental results; since processing a meaning in situation needed longer time than that on word meaning. However, Storto and Tanenhaus (2005) found that the processing of SI is quite local and the generation of it was very rapid and could be completed below the level of a whole sentence comprehension. Huang and Snedcker (2009) also provided evidence that SI can be processed online locally. Both of the two studies show disagreement to the context driven account.

1.3.2 Napoleon Katsos' Study

Napoleon Katsos participated in the debates and conducted a few relevant experiments for a preferred model. Katsos and his colleague, Dorothy *Bishop,* conducted a study in 2009 about the processing model of the scalar implicature. They adopted both the classic scale (some < all) and the poset scale (triangle<{triangle, square}) to distinguish the context driven and the default account. In their experiments, they compared SI acquisition of the two different scales between adults and children. They employed the experimental mechanism of Truth Value Judgment Task, and the experimenter needs to judge the utterance made by a cartoon character named Mr. Caveman, who watched a situation with the experimenter and then used one sentence to describe it. If participants were able to make correct judgments about Mr. Caveman's answer, they might have recognized the existence of SI. According to the final

result, Katsos and Bishop found the unreliability of default account but a richer picture than what the context-driven account expects. In Katsos (2011), he and Dorothy *Bishop* conducted another study based on the one in 2009, and similar results were conducted.

1.3.3 Garrett and Harnish's study

The most representative study that holds the standardization model was the one conducted by Garrett & Harnish (2007). In their experiments, Garrett & Harnish tested a particular type of GCI, (which was termed *implic**i**ture* by Bach). They made a clear distinction concerning a psychological reality of both the default account, especially Levinson's I-heuristic, and Bach & Harnish's standardization account. They provided two kinds of contexts, the enabling context (supporting the process of *impliciture*) and the canceling context (obstructing the process of *impliciture*). Both types of contexts contained the same target sentence in them. Following each of the contexts, a question concerning comprehending of a target sentence was presented. The reading time of the target sentence and the vocal answer of the question were recorded. Through a technical statistical analyses of the data, they concluded that the standardization model was more preferred than default model, because the experimental results consistently favored both the enabling and the canceling contexts, and the difference in reaction time was not significant under both contexts. This result essentially conflicted with the default and the context-driven accounts, because the default account presumes that the GCI would be first canceled and then processed in the canceling context (i.e. the lower bound context) and thus with a longer reaction time, while in the context driven account, the GCI would not be processed at all in the canceling context.

Based on the previous experiments mentioned above and in order to seek out more solid evidence for further investigation, we designed our experiment as follows.

1) We focus on the classic scale <some, all>， hoping to obtain clear and reliable results about the processing model of scalar implicature through a comparison between the processing of the children and the adults. As to the Chinese material, we adopted "*yixie*" as the equivalent to "some". This choice is based on the research conducted by Wang Chunmei (2014). Three Chinese words, like "*youde/youxie/yixie*" owning the meaning of "part of a whole", were tested their appropriateness used in the sentences with scalar implicature by using Likert scale questionnaires to 30 Chinese native speakers. The result showed that 58% of participants expressed much more inclination to the "*yixie*". In this way, we adopted "*yixie*" as the equivalent to "some" in our testing materials.

2) In order to better suit the experiment of children group, we adopted a similar experiment form as in the Katsos (2009) and Katsos (2011). All the testing materials are in the form of picture and include a character help to conduct the experiments. The difference lies in the role of this cartoon character, which is the result of the change of experimental paradigm. The paradigm of the Truth Value Judgment Task, which was used in Katsos's experiment, contained a judgment to the expression of a third person according to the situation shown, instead of the direct conversation between the addresser and the addressee. Thus, the whole processing in the experiment might have been unnatural and indirect, which might have led to a somewhat different reaction and process from the real situation in the daily conversation.

We changed the experiment paradigm, the TVJT, into the participant- perception test, which asked the participant to directly react to the questions raised by the cartoon character instead making judgments.

3) In order to seek out the theoretical reason for the psychological priority revealed in the scalar implicature processing of classic scale, namely, whether it is contributed to the context driven account, the default account or the standardization, we determined to first cut off the context and merely conduct an experiment with only the stimulus of literal form (The hypotheses of the distinctions among the three accounts is in detail shown in Table 1). Then we added the context to the experiment mentioned above to examine the difference in results. In both experiments, we combined the other experimental factors like the reaction time together with the correctness rate of the participants in order to provide more evidence for the determination of the theoretical hypothesis.

3. Methods in Experiment 1

3.1 Objectives

Experiment 1 was an online experiment. The purpose of this experiment was to distinguish the three major models, more specifically, to test the necessity of presence of context by cutting off contexts in the testing materials.

3.2 Hypothesis

We presumed that in the context driven account, the scalar implicature would not be processed in the situation without any contexts in both child and adult groups. In the default account, the correctness rate and reaction time of child group would be at the similar level with the adult group, due to the default mechanism possibly happening in both child and adult processing. In the standardization account, though the literature had not provided any information about definite reaction for children to perform the standardization in their mental capacity, we made a prediction about a possible weaker capacity than or similar capacity to the adult, and a longer reaction time cost in children's scalar implicature processing. Compared with the adults, the children experienced much less exposure to scalar implicature in their daily conversation; hence, for the child group, we predicted a possible similar trend of SI correctness rate accompanied with a much slower reaction time compared with the adults (See Table 1).

Table 1. The hypothesis of Experiment 1

	PR		RT	
	adults	children	adults	children
Context-driven	Not processed	Not processed	-	-
Default account	high	high	quick	quick
Standardization	high	high	quick	slow

3.3 Participants

The participants consisted of an adult group and a child group. Twenty adults who were 20 to 25 years old were randomly chosen from Lanzhou University. In order to control the possible

influence from the discipline students are in, we intentionally selected half of the participants from the humanity major and the other half from the scientific major. The numbers of the male and the female are equally divided based on the total amount of participants. Eighteen children were randomly chosen from a preschool class at the a kindergarten in Lanzhou, China. All the participants had no knowledge of linguistics and related fields.

3.4 Materials

We used the scale <some, all> without context in this experiment. Both the picture and the vocal demand were presented according to the participant perception-test. In each picture, there were several objects of two or three kinds and a basket at the right bottom. After hearing the vocal demands ("Please put ____ into the basket" in Chinese), the participant needed to speak out number they wanted to put into the basket. In this way, they can give the direct reactions to the scalar implicature themselves. The details of the testing materials are as follows.

As to the picture, there were three types, two for the testing group and the other one for the controlling group. The amounts of each kind of objects in all pictures were different. The key difference between these groups lay in the vocal demand. The model of the vocal demand was "Please put ____ into the basket" ("qing wang lanzili fang ru _____" in Chinese version). In the testing group, the vocal demand was "Please put some____ into the basket" ("qing wang lanzili fang ru yixie _____"). In the Control Group A, the vocal demand was "Please put all____ into the basket" ("qing wang lanzili fang ru suoyou _____"). In the Control Group B, the vocal demand was "Please put (specific number of)____ into the basket" ("qing wang lanzili fang ru __ ge _____"). There were two points of the design intention needed to be stated. First, the purpose of the design with different number of two or three objects in the same picture was to avoid the stylized answering. Second, in the Chinese version of vocal demand, the target object of each utterance was put at the end of the sentence ("qing wang lanzili fang ru _____"). In this way, the participants had to finish listening to the vocal demand before giving the answer. Beisdes, to put the target item at the end of the utterance also made it easy to record the reaction time of the participants. Because all the vocal demands were of the same length (also with the same amount of characters), the only difference lied in the target object at the end of each sentence. In this way, the difference of reaction time in each sentence can only be caused by the processing of the target testing items, which made the recording and analysis of the reaction time easier.

Before the real test, both adult and child participants received training. The purpose of the training was to help participants become familiar with the form of the testing material, especially speaking and pressing the right Shift button at the same time. The content of training comprised of two parts. The first part let the participants count out how many objects were on the screen. A picture contained 5 to12 objects of the same kind that were arranged in several lines with 5 objects on each of the lines (See Figure 4). The second training was similar to and in the same form as the real test (See Figure 1–3). Some examples of picture materials are as follows (all the pictures used in Figure 1-4 are from the website: http://image.baidu. com/):

Testing group (See Figure 1)

qing wang lanzi li fangru yixie pingguo.

please to basket in put some apples

Please put some apples into the basket.

Figure 1. An example of the material used in testing group

Control Group A (See Figure 2)

qing wang lanzi li fangru suoyou pingguo.

please to basket in put all apples

Please put all the apples into the basket.

Figure 2. An example of the material used in control group A

Control group B (See Figure 3)

qing wang lanzi li fangru liang ge qiezi.

please to basket in put 2 (quantifier) eggplant

Please put 2 eggplants into the basket.

Figure 3. An example of the material used in control group B

Figure 4. Other examples of the material used in Part 1 of training

3.5 Procedures

The adult groups were processed by a tester in the sound-proof rooms in the Experimental Pragmatics Laboratory of our college. The children groups were processed in a quiet room in their kindergarten. All the training and the tests are presented through the software of DMDX on the computer. The whole processes of training and tests were finished on the individual basis. The detailed procedure is as follows.

Two types of training to the participants went first.

1) In the first training, the participants were asked to count out the number of the same kind of object shown on the screen. The participants were asked to speak aloud the number of the objects in the picture as quickly as they could and to press the right Shift button at the same time. The material is shown in Figure 4.

2) In the second training, the participants were asked to finish the same task as that in the formal test, whose material is shown in Figure 1-3. The participants were asked to look at the picture and listen to the vocal demand first and then spoke out the number of objects they would like to put into the basket. They did the same operation to the button as in the first training.

Then, the formal test ran as in the following steps.

1) The participants were shown a picture presented in which there were several objects of two or three kinds. On the bottom right of the picture, there was a basket.

2) The participant heard a voice saying, “Please put... into the basket”.

3) The participant was asked to respond with how many objects he or she was going to put in the basket in order to meet the vocal demand, and at the same time pressed the right Shift button to record the reaction time.

4) The vocal answers and reaction time were recorded by the DMDX.

3.6 Results

The results we obtained from Experiment 1 mainly focused on the correctess rate and the reaction time. We counted the answer between zero and the total number of the target object as the right answer, while the answer with zero or the sum and the no-answering as the wrong answer. The details of the results were as follows.

1) The correctness rate of testing items with “some” of the adult group was 62%, which was not significantly different ($p=0.102 > 0.05$) from the rate of the child group, 59%. The correctness rates of testing items “all” are 43% and 97% in the child and adult groups

respectively, with significant difference (p=0.000<0.05). The correctness rates of testing items with specific number are 70 % and 99% in the child and adult groups respectively, with significant difference (p=0.000<0.05). (See Table 2 and Table 3).

Table 2. The correctness rate of both the child and the adult groups

Correctness rate	subjects	N	Mean	SD	SEM
some	children	302	.59	.493	.028
	adults	320	.62	.486	.027
all	children	152	.43	.496	.040
	adults	157	.97	.158	.013
specific number	children	152	.70	.461	.037
	adults	159	.99	.079	.006

Table 3. Independent Sample Test for the scalar implicature correctness rate of both the children group and the adult group

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. 2-tailed	MD	SED	95% Confidence Interval of the Difference	
									Lower	Upper
Equal variance assumed	some	2.682	.102	-.831	620	.406	-.033	.039	-.110	.045
	all	1069.247	.000	-13.135	307	.000	-.547	.042	-.629	-.465
	specific number	675.091	.000	-7.985	309	.000	-.296	.037	-.369	-.223

2) The reaction time of the adult group to “some” was 599.06 ms, which was much faster than that of the child group (1731.36 ms) with significant difference (p=0.00 < 0.05). As to the items with “all”, the reacion time of the adult group was 505.51 ms, which is much faster than the child group (1847.95) with significant difference (p=0.00 < 0.05). The reaction time of specific number are 344.96 ms and 1425.21 ms in the adult and the child group respectively, which shared significant difference (p=0.00 < 0.05). (See Table 4 and Table 5).

Table 4. The reaction time of both the child group and the adult group

Reaction Time	subjects	N	Mean	SD	SEM
some	children	254	1731.36	942.001	59.106
	adults	296	599.06	507.334	29.488
all	children	108	1847.95	947.625	91.185
	adults	144	505.51	414.663	34.555
specific number	children	134	1425.21	798.638	68.992
	adults	146	344.96	239.251	19.801

3.7 Analysis and Discussion

First, according to the first result mentioned above, though without context, both the adults and children still process scalar implicature. In other words, most of the participants can

process the scalar implicature merely through the literal form of scalar items even without the

Table 5. Independent sample test of the reaction time between the child and the adult groups

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	MD	SED	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	some	108.524	.000	17.880	548	1132.308	63.329	1007.910	1256.705
	all	94.712	.000	15.179	250	1342.440	88.439	1168.260	1516.620
	specific number	118.858	.000	15.601	278	1080.253	69.243	943.946	1216.559

support of context. This result seems to be in contrast with the hypothesis of context-driven account, and provides possible evidence for the default account and the standardization account.

Second, according to the second result, the reaction time of children in the test items with "some" is much slower than that of the adults. This result might question the presumed similar reaction between the child group and the adult group held by the default account, which is due to the possible same work of the default mechanism in both participant group. The slower reaction time of the child group can be explained by the less accumulation of standardized scalar implicature in the child group than to the adult group.

In this way, from the results of Experiment 1, we find the certain evidence to deny the possibility of the context account and raise question about the default account. The results are seem to be much closer to the hypothesis of the standardization.

However, in the after-experiment interview, about 30% of the adult participants responded that they did not realize any difference of using "some" and "all" in the testing scenarios, although they were aware that the affirmation of "some" meant the denial of "all". For example, when the vocal instruction was, "Put some of the rabbits into the basket", the participants did not feel that it was inappropriate to put all 10 rabbits into the basket, since no specific demand in the experimental context asked for a distinction between "some" and "all" and all 10 of the rabbits would be "some" considering in a wider and more abstract notion of rabbit if there is no obvious constraints provided. It was a little confusing for them when giving reaction to the testing material without any context constraints, especially when being asked to speak out the specific amount of objects to put in the basket. That might be accompanied with various unnecessary guesses about the testing purpose and give responses according to their personal predictions. It would be necessary for us to get rid of such guesses from the participants before making the final conclusion.

Our solution is to add context constraints, which can not only help participants decide the specific amount of objects, but also might be further helpful in making a distinguish between

the default account and the standardization account. It is necessary to point out that though our present results of this context-free experiment have testifies the reliability of both the default and standardization account, it does not mean that these two theories propose the processing of SI utterance can only happen without contexts and thus context is of no significance. In fact, one of the key distinctions between the default account and the standardization lies in their different hypothesis of SI reaction in two different contexts, which is the upper bound context which supports the derivation of scalar implicature, and the lower bound context that interfere the SI utterance processing. The default account holds that the processing of utterance with SI in the lower bound context is much consuming because the cancellation might happen, which is revealed with the longer reaction time in the lower bound context. The standardization account does not take a difference in the process of scalar implicature in both the upper bound and the lower bound contexts. The implicature would be retained even if the context is not suitable, for the implicature has already been standardized. In this way, it would be better to add the upper bound context and the lower bound context at the same time, which might serve as another convincing element for distinguishing the default and the standardization accounts.

4. Methods in Experiment 2

4.1 Objectives

As has been mentioned above, the main purpose of Experiment 2 is to distinguish the theories of the default account and the standardization account on the basis of avoiding the possible guesses from the participants with the help of adding contextual constraints.

4.2 Hypothesis

The presumption of this experiment is as follows. On the standardization account, the scalar implicature has been fixed on the literal form of the scalar items, and thus the scalar implicature would still be processed even when the context is not appropriate. The reaction time of processing scalar implicature in the upper bound context and the lower bound context would be no significantly different. On the default account, due to the stereotypical information would be triggered through some specific lexical forms and relevant discourse topics, the processing of utterance with SI would be abstained in the lower bound context. Besides, the reaction time would be longer in the lower bound context due to the possible cancellation of the previously lexically triggered scalar implicature after encountering the less supportable factors in the lower context.

4.3 Participants

We merely adopt the adult group as the participants in the Experiment 3, for there is no clear literature theory which would provide a definite hypothesis to the children reaction under the framework of the standardization. Besides, this is out of the consideration of the special characteristic of scalar implicature standardization. The adults are much highly exposed to the scalar implicature in their daily conversation and thus they would provide a more stable and certain reaction especially towards the standardization of scalar implicature. In this way, through this reliable reaction of the adults, we might reach a clearer result to distinguish the

default account and the standardization account. All of our participants are students form Lanzhou University, who are aged from 20-25. The selective criteria are the same as that in the Experiment 1.

4.4 Materials

We continuously adopted the scale of <some, all> and the testing mechanism of the participant perception test. Besides, as has been mentioned above, in this experiment, we add the context constraint of the upper bound and the lower bound contexts in order to control the unnecessary guesses from the participants.

The testing materials were divided into two testing groups (the upper bound and the lower bound respectively) and one control group. An example of the materials of the testing group and the control group is as follows.

Testing items

Upper bound context

Scenario:

ni lai xiaodian li mai pingguo
you come small store in buy apples.
You come into a small store to buy apples.

xianzai gong you 8 ge pingguo chushou
now totally there be 8 (quantifier) apples for sale
There are totally 8 apples for sale now.

jintian pingguo youdian gui, ni zhi xiang mai 3 dao 4 ge.
today apple a bit expensive, you only want buy 3 to 4 (quantifier)
The apple today is a bit expensive. You only want to buy 3 to 4 apples.

Instruction:

qing wang nide gouwu lan li fangru yixie pingguo.
please to your shopping basket put in some apples
Please put some apples into your shopping basket (for check-out).

Lower bound context

Scenario:

ni lai xiaodian li mai pingguo
you come small store in buy apples.
You come into a small store to buy apples.

xianzai gong you 8 ge pingguo chushou
now totally there be 8 (quantifier) apples for sale

There are 8 apples for sale now.

jin tian jia li lai le keren, ni xuyao mai 8 ge.
today home at come (particle) guests, you need buy 8 (quantifier)
Guests are visiting your home today. You need to buy 8 apples.

Instruction:

qing wang nide gouwu lan li fangru yixie pingguo.
please to your shopping basket put in some apples
Please put some apples into your shopping basket (for check-out).

Control items

Scenario:

ni lai xiaodian li mai pingguo
you come small store in buy apples.
You come into a small store to buy apples.

xianzai gong you 8 ge pingguo chushou
now totally there be 8 (quantifier) apples for sale
There are 8 apples for sale now.

zhe xie bu shi ni xihuan chide ruan pingguo, ni xiang mai 3 ge.
these not be you like eat soft apple you want buy 3 (quantifier)
These are not the soft apple you like to eat. You want to buy 3 apples.

Instruction:

qing wang nide gouwu lan li fangru sange pingguo.
please to your shopping basket put in 3 (quantifier) apples
Please put 3 apples into your shopping basket (for check-out).

The detailed explanation of the contexts is as follows. In the material about the testing items, we add a situation in which the participants need to act as the customer of a small shop. This setting is shown in the first sentence. The second sentence illustrates a notion of "all" to the participants, which can provide the participants a better understanding of the context. The most important testing point is at in the third sentence, which shows differences in the upper bound context and the lower bound context. In the upper bound context, the third sentence provides a situation in which the processing of utterance with SI would be more appropriate. More specifically, the upper bound context provides some constraint situations, like the object itself does not meet the demand described in the situation, or it does not need to buy too many objects according to the presented scenario. In this way, the scalar implicature of "some" implies "not all" would be more acceptable. Besides, we reinforce the possible "not all" response through adding an intention of buying specific amount of objects. Thus, this kind of upper bound context would further enable the processing of utterance with SI.

However, in the lower bound context, the third sentence might disturb the processing of utterance with SI. It shows a situation in which all the objects are wanted, like having lots of guests at home or showing a strong personal preference to the objects. In order to make a strong distinction with the upper bound context and also set a contrast to the notion of "some", we add the number of all the objects as the possible buying intention in the third sentence. Thus, "some" might imply "possibly all" and the scalar implicature would be inappropriate in this context. The last sentence is the same in both the upper bound and the lower contexts, which is the demand the participants need to finish according to the situations mentioned in the previous context.

In the material about the control items, the pattern of the context is similar to the testing ones. The only difference lies in the last sentence of demand, in which the specific amount of objects put in the basket is directly given to the participants.

All the materials, including the testing items and the control items, are shown in the form of animation through the software of DMDX on the computer. There are vocal explanations of the animation in each settings, whose contents are exact the same as the literal materials mentioned above. An example of the material in the picture form is shown as follows. The order of picture in each context is in the same order as the order of the literal material shown above.

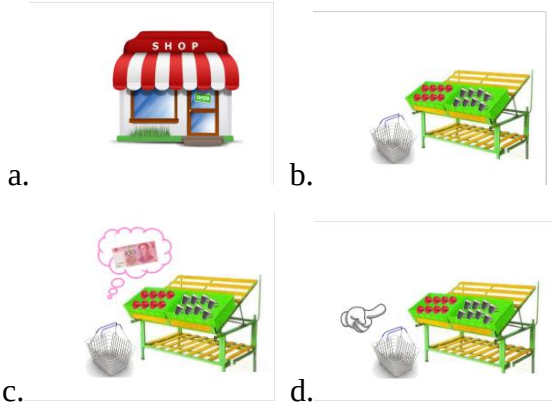

Figure 5. Testing materials of the upper bound context in Experiment 2 where (a) indicates the shop, (b) the items sold in the shop, (c) the target items are a bit expensive, and (d) the instruction to put some of the target items into the basket

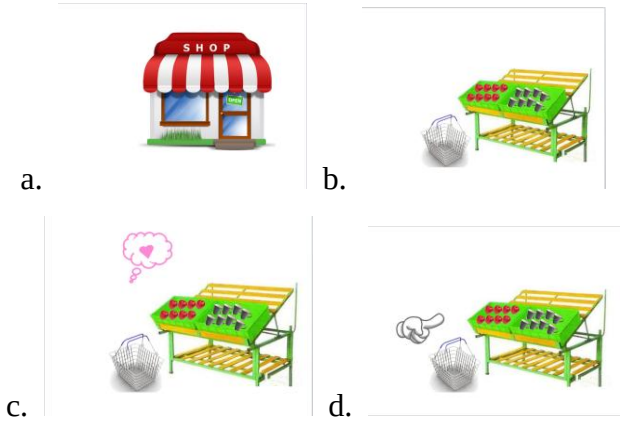

Figure 6. Testing materials of the lower bound context in Experiment 2 where (a) indicates the shop, (b) the items sold in the shop, (c) the target items are all favored under specific conditions, and (d) the demand for putting some of the taege items into the basket.

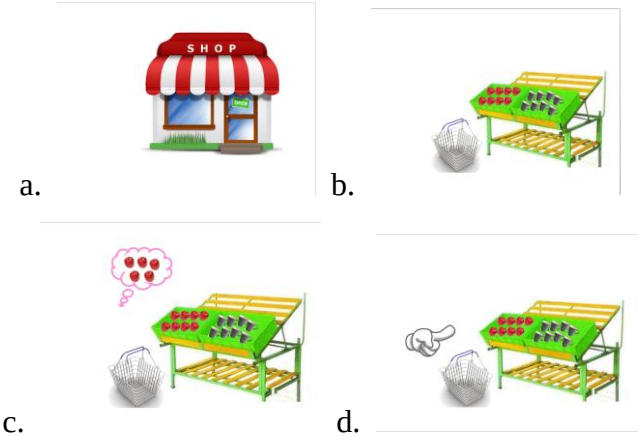

Figure 7. Testing materials of the control items in Experiment 2 where (a) indicates the shop, (b) the items sold in the shop, (c) a specific amount of the target items are all favored, and (d) the demand for putting some of the target items into the basket.

4.5 Procedures

The testing materials were shown on the computer through the software of DMDX. The whole processes of training and tests were finished on the individual basis. The details of each procedure in the experiment is as follows.

1) The participants were asked to watch the scenario composed of several pictures and the vocal explanation.

2) The participants heard the vocal demand of "qing wang nide gouwulanli fangru yixie_____." (Please put some ….. in the basket for check-out.)

3) Participants were asked to speak out the specific amount of objects they determined to put into the basket according to their understanding of the scalar item "some" and the specific context. They were also asked to press the right shift button at the same time of response,

4) The vocal response would recorded by the software DMDX. The reaction time also be

recorded through the pressing of the right shift button by DMDX.

The whole procedure is accomplished in the sound proof room in the Experimental Pragmatic Lab of Lanzhou University, China.

4.6 Results

The testing results of the Experiment 2 were also consisted of both the correctness rate and the reaction time. As in the Experiment 1, the answer with number between zero and the sum of the target object was counted as right, while the answer with zero and the total number and the no-answering as wrong. The results are as follows.

1) In the aspect of correctness rate, the testing result shows that the correctness rate of scalar implicature in the upper bound context is higher than that in the lower bound context (See Table 6), which shows a significant difference between the two contexts. (See Table 7)

2) In the aspect of reaction time, the processing of the scalar implicature in the upper bound context and the lower bound context shows close speed, though the processing in the upper bound context is slightly quicker than the lower bound context (See Table 6). There is also not significantly different between the reaction time in both contexts (See Table 8).

Table 6. The mean of correctness rate and the reaction time

	correctness rate	Reaction time
Upper bound context	94%	1506.39
Lower bound context	51%	1393.36
Control group	97%	1322.40

Table 7. The result of the paired sample test of the correctness rate

Mean	SD	SEM	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
			Lower	Upper			
.475	.929	.074	.329	.621	6.425	157	.000

Table 8. The result of the paired sample test of the reaction time

Mean	SD	SEM	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
			Lower	Upper			
-158.366	1254.547	99.181	-354.248	37.515	-1.597	159	.112

4.7 Analysis and Discussion

According to the testing results of the Experiment 2 mentioned in the previous section, we find out that the model shown in the results is much closer to the standardization account. The detailed analysis is as follows.

1) The first point of the result reveals a high correctness rate in the upper bound context and a medium correctness rate in the lower bound context. First, this result shows a retained process even in the lower bound context, in which the processing of utterance with SI is not favorable. Second, in the lower bound context, in which the intention of purchasing all the objects has been clearly given through providing the specific number of all the objects, there is still 51% of correctness rate of scalar implicature under such a kind of circumstance. It seems that the strong influence from the lower bound context might not fully determine the processing of utterance with SI. This falls exactly into the theoretical framework of the standardization account, which holds the assumption that the scalar implicature has already been standardized or fixed in the literal form. Thus, the context may not be decisive in the scalar implicature processing and the scalar implicature would be processed in both the enabling context (the upper bound context) and the canceling context (the lower bound context),, which is in accordance with the results in the experiments of Garrett & Harnish (2007). In contrast, the default account persists that the lower bound context would cause a much greater impact on the processing of utterance with SI. The reason is that, according to the default account, the stereotypical information needed to process scalar implicature is partially determined by the background information not only like the specific lexical items but also the evolving topics of the current context. Such related discourse topics in the lower bound context is usually in a disturbance with the scalar implicature.

2) The second point of the results shows no significantly different reaction time between the upper bound context and the lower bound context, which also testifies the presumption of the standardization account and the results of Garrett & Harnish (2007). As has been mentioned above, the standardization account presumes that the scalar implicature has already been fixed through the long time accumulation in the daily conversation. Thus the reaction time towards such a fixed notion of scalar implicature would be quite close in any contexts. However, the default account presumes a possible cancellation of scalar implicature in the lower bound context during the processing, which would cause an observable delay in the reaction time of scalar implicature in the lower bound context.

5. Conclusion

In this study, we mainly focus on finding evidence for the processing model of scalar implicature. In Experiment 1, we adopted the scale <some, all> as the testing material without context and recorded the reaction time of each testing items. We found that both the child and the adult group can process scalar implicature even under the situation without any context, which can serve as a solid evidence for the denial of the context driven account. Besides, there was no processing priority in children group, which was in contrast with the default account. However, we still needed to overcome the unnecessary guess from the participants about the experimenter's intention due to the lack of certain contextual constraints. In Experiment 2, we added the upper bound context and the lower bound context to the same material, and we adopted only the adult participants in this experiment. Through testing and statistical analysis, the final result showed a retained processing of utterance with SI in the lower bound context and a not significantly different reaction time between the upper bound and the lower bound contexts. Such a kind of result shares a closer relationship

with the hypothesis of standardization.

The statistics from our experiments showed an unfavorable result to the context driven account and the default account. Though the preliminary results of our Experiment 2 was in accordance with the standardization account, further evidence is still needed to be found to get a closer and detailed examination about this theory. Besides, the study on the children reaction and acquisition to scalar implicature standardization is also a necessity, which might provide a more comprehensive understanding to the theory of standardization.

Acknowledgements

We authors owe a great deal of gratitude to the sponsors. This work was sponsored by the 2011 Humanities and Social Science Planning Fund of the China Ministry of Education under Grant [number 11YJA740053]. It is also supported by the Constructive Program Funds for Teaching and Research Innovation Teams (Project No. 16LZUWYXSTD017) of the School of Foreign Languages and Literature, Lanzhou University.

References

Bach, Kent (1994). Conversational impliciture, *Mind and Language*, *9*, 124-162.

Bach, Kent (2000). Quantification, qualification, and context: A reply to Stanley and Szabo, *Mind and Language, 15*, 262-283.

Bach, Kent (2005). Context ex machina, in Z. Szabó (ed.), *Semantics vs. Pragmatics*, Oxford: Oxford University Press, pp. 15-44.

Bach, Kent (2006b). The excluded middle: Semantic minimalism without minimal propositions, *Philosophy and Phenomenological Research, 73,* 435-442.

Bach, Kent (2010). Context Dependence (such as it is), *The Continuum Companion to the Philosophy of Language.*

Bezuidenhout, A. & Cutting, J.C. (2002). Literal meaning, minimal propositions, and pragmatic processing. *Journal of Pragmatics, 34,* 433-456.

Bezuidenhout, Anne (2002). Truth-conditional pragmatics. Philosophical Perspectives, *16,* 105-134.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*, 437-457.

Breheny, R., Katsos, N., & Williams, J. (2006). Are Generalized Scalar Implicatures Generated by Default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*(3), 434-463.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and Implications*. Amsterdam: John Benjamins.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*,

Blackwell, Oxford.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatic interface. In Adriana Belletti (ed.), *Structures and Beyond. The Cartography of Syntactic Structures, Volume* 3, *Oxford Studies in Comparative Syntax*, pp. 39–103. Oxford, New York: Oxford University Press.

Garret, M., & Harnish, R. M. Experimental pragmatics: testing for implicitures. *Pragmatics and Cognition, 15*(1), 65-90.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo F., Gualmini, A., & Meroni L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes, 20*(5), 667-696.

Hirschberg, J. (1991). *A Theory of Scalar Implicature*, Garland, New York.

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental psychology, 45*(6), 17-23.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development, 2*, 77-96.

Katsos, N., & Bishop, D. V. M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature, *Cognition*, 2011.

Katsos, N., & Cummins, C. (2012). Scalar implicature: Theory, processing and acquisition， *Nouveaux cahiers de linguistique fran ̧aise, 30*(2012), 39-52.

Katsos, N., & Cummins. C. (2010). Pragmatics: From theory to experiment and back again, *Language and Linguistics Compass, 4*(5), 282-295.

Katsos, N. (2007). The Semantics/Pragmatics Interface from an Experimental Perspective: the Case of Scalar Implicature, *Synthese, 165*(3), 385 - 401.

Katsos, N. (2009), Evaluating under-informative utterances with context-dependent and contextindependent scales: experimental and theoretical implications, In Sauerland, U. & Yatsushiro, K., *Experimental Semantics and Pragmatics*, 51-73.

Levinson, S. (2000). *Presumptive Meanings*. Cambridge, MA.: MIT Press.

Liu S., & Liu Huangmei (2014). A Review of Cognitive Model: Experiments of Scalar Implicature. *Studies in Literature and Language, 8*(3), 22-28.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language, 85*(2), 203-210.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: experiments at the semantics/pragmatics interface. *Cognition, 86*, 253-282.

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language*

Acquisition, 12, 71-82.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Storto, G., & Tanenhaus, M. K. (2005). Are scalar implicatures computed online? *Proceedings of SuB, 9,* 431-445.

Zhao, M., Liu, T., Chen, G., & Chen, F. (2015). Are scalar implicatures automatically processed and different for each individual? A mismatch negativity (MMN) study. *Brain research, 1599,* 137–149.

Zondervan, A. (2010). *Scalar implicatures or focus: An experimental approach.* Unpublished doctoral dissertation, Universiteit Utrecht, Amsterdam.

Copyrights